

(12) UK Patent Application (19) GB (11) 2 359 153 (13) A

(43) Date of A Publication 15.08.2001

(21) Application No 0002909.0

(22) Date of Filing 10.02.2000

(71) Applicant(s)
International Computers Limited
(Incorporated in the United Kingdom)
26 Finsbury Square, LONDON, EC2A 1SL,
United Kingdom

(72) Inventor(s)
Trevor Hall

(74) Agent and/or Address for Service
D C Guyatt
International Computers Limited, Intellectual
Property Dept, Cavendish Road, STEVENAGE,
Hertfordshire, SG1 2DY, United Kingdom

(51) INT CL⁷
G06F 9/46

(52) UK CL (Edition S)
G4A AFL

(56) Documents Cited
<http://sup.datadirectnet.com/drivers/emulex/sli124.txt>
14.12.99, Emulex Corporation

(58) Field of Search
UK CL (Edition R) G4A AFL
INT CL⁷ G06F 9/46 13/40 13/42
ONLINE: COMPUTER EPODOC JAPIO WPI INTERNET

(54) Abstract Title
Device discovery in a shared data storage system

(57) A computer system using the SCSI protocol has at least two processing nodes and a shared data storage system comprising a number of storage units. Each storage unit has a logical unit number (LUN). A first processing node on boot-up performs a device discovery procedure in which it scans the shared data storage system to find storage units with LUNs up to a maximum LUN value stored in non-volatile storage in the node. A second processing node on boot-up performs a device discovery procedure in which it scans the shared data storage system to find storage units with LUNs not less than a start LUN value stored in non-volatile storage in the node, and up to either a stop value stored in the node, or up to the start value plus a constant.

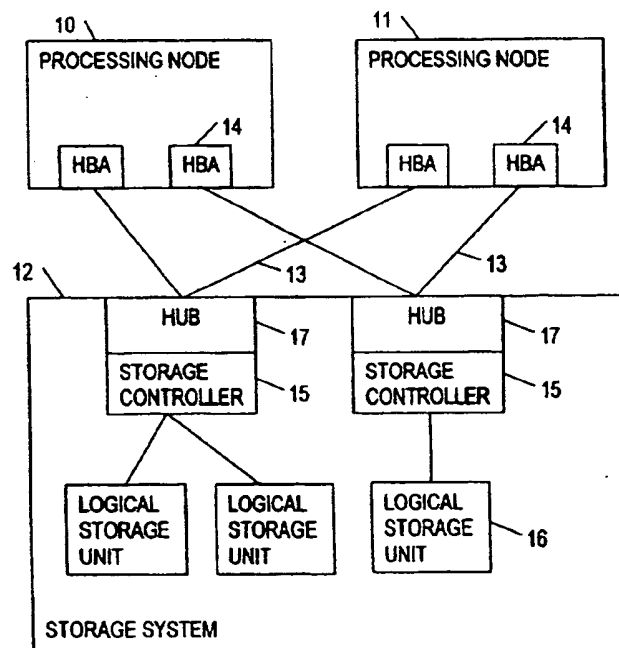
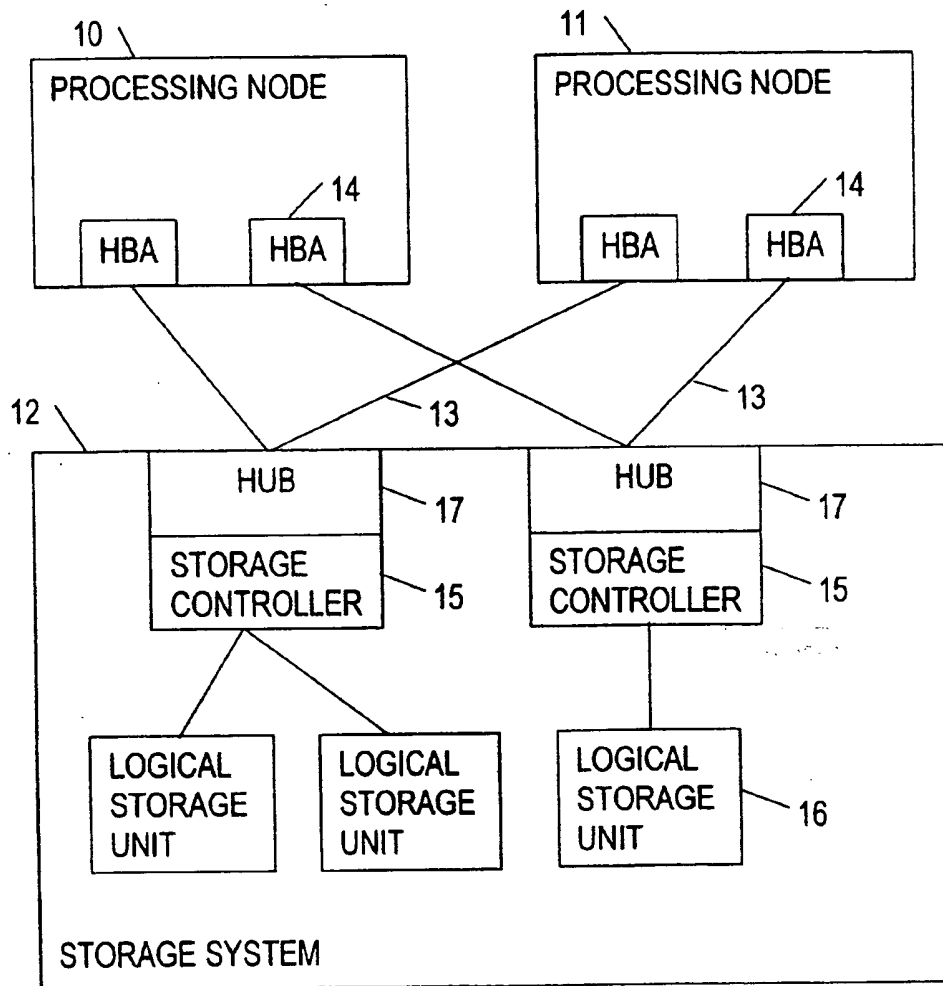


FIG.1

GB 2 359 153 A

**FIG.1**

2/2

```
READ MAXLUN FROM NVRAM
FOR EACH HBA
  FOR EACH TARGET DEVICE
    SEND TEST UNIT READY
    IF VALID RESPONSE RECEIVED
      FOR LUN=0 TO MAXLUN
        SEND INQUIRY
        SAVE RETURNED DATA
      NEXT LUN
    END IF
  NEXT TARGET DEVICE
NEXT HBA
```

FIG. 2

```
FOR EACH HBA
  READ STARTLUN FROM CONFIGURATION FILE
  FOR EACH TARGET DEVICE
    SEND TEST UNIT READY
    IF VALID RESPONSE RECEIVED
      FOR LUN=STARTLUN TO STARTLUN + 31
        SEND INQUIRY
        SAVE RETURNED DATA
      NEXT LUN
    END IF
  NEXT TARGET DEVICE
NEXT HBA
```

FIG. 3

MASKING LOGICAL UNIT NUMBERS IN A SHARED DATA STORAGE SYSTEMBackground to the Invention

This invention relates to a method and apparatus for masking logical unit numbers in a shared data storage system.

A data storage system typically consists of an array of disk units and one or more storage controllers. The disks may be accessed using the well-known SCSI protocol, over a Fibre Channel (FC) connection. Within a SCSI storage system, groups of disks are bound into logical storage units, each of which is assigned a logical unit number (LUN). Under SCSI, a storage unit is addressed by a SCSI target address, which selects a particular storage controller, and by the LUN, which selects one of the storage units connected to that controller. For a stand-alone disk, the LUN is always zero.

Such a storage system may be shared between two or more processing nodes, possibly running under different operating systems. In such a case, it is necessary to provide some form of LUN masking, to ensure that each node can access only its own LUNs, and cannot access LUNs belonging to another node.

There are two approaches currently being taken to LUN masking. In a first approach, LUN masking is performed within each storage controller, using the Worldwide Name (WWN) of the FC Host Bus Adapter (HBA) of each processing node to map LUNs on to the nodes. In a second approach, LUN masking is performed within each processing node, using the unique WWN of the storage controller. Masked LUNs are hidden from the operating system.

Both of these known LUN masking techniques require special software, running either within the storage system, or in the processing nodes. Such software is generally expensive.

The object of the invention is to provide a novel solution to the problem of LUN masking, which is less expensive than known solutions.

Summary of the Invention

According to the invention a computer system comprises:

- at least two processing nodes; and
- a shared data storage system comprising a number of storage units, each having a logical unit number (LUN) associated with it;
- wherein a first of the processing nodes stores a maximum LUN value in non-volatile storage, and is arranged to perform, on boot-up, a device discovery procedure in which it scans the shared data storage system to find storage units with LUNs not greater than the maximum LUN value;
- and wherein a second of the processing nodes stores a start LUN value in non-volatile storage, and is arranged to perform, on boot-up, a device discovery procedure in which it scans the shared data storage system to find storage units with LUNs not less than the start LUN value.

It will be shown that the invention allows LUN masking to be achieved relatively inexpensively with only minimal modification to the operating systems of the processing nodes, using existing device discovery features of the operating system.

Brief Description of the Drawings

Figure 1 is a block diagram of a computer system, comprising a number of processing nodes and shared storage.

Figure 2 illustrates a device discovery operation performed at boot-up in one of the processing nodes.

Figure 3 illustrates a device discovery operation performed at boot-up in another of the processing nodes.

Description of an Embodiment of the Invention

One LUN masking technique in accordance with the invention will now be described by way of example with reference to the accompanying drawings.

Figure 1 shows a computer system comprising a number of processing nodes 10, 11 and a shared data storage system 12. The storage system is connected to the processing nodes by way of Fibre Channel (FC) loops 13, using the FCP SCSI-3 protocol.

Each of the processing nodes 10, 11 runs a particular operating system. In this example, node 10 runs the ICL VME operating system, while node 11 runs either the Windows NT operating system, or the Unix operating system. [Windows NT is a registered trademark of Microsoft Corporation in the USA and/or other countries. Unix is a registered trademark of The Open Group]. Each processing node includes one or more FC Host Bus Adaptors (HBAs) 14, for connecting the node to the Fibre Channel loops.

In the present example, the storage system 12 is an EMC CLARiiON C5700 Storage Array. It includes two storage controllers 15, with built in two port hubs 17. Each controller has a number of logical storage units 16 associated with it. Each logical storage unit 16 typically comprises a RAID group of disks.

Each storage controller 15 has a Fibre Channel loop identity (ID). Each logical storage unit 16 is assigned a logical unit number (LUN) which, in this case, is unique within the storage system. A logical storage unit is addressed by the combination of the loop ID of its storage controller, and its LUN.

At loop initialization time, the loop IDs of all devices (HBAs, storage controllers) connected to each FC loop are circulated around the loop. These IDs are stored within each HBA. From the Fibre Channel process login parameters, it is possible to determine which devices are initiators (in this case, the HBAs), and which are targets (in this case, the storage controllers).

At boot-up time, each of the processing nodes 10, 11 performs a device discovery (or "grope") procedure, in which it attempts to determine which devices are connected to the node.

Figure 2 is a pseudo-code representation of the grope procedure for the Windows NT or Unix node 11. It should be noted that this grope procedure is conventional for Windows NT or Unix systems, and hence does not require any special software or modification to the operating system.

The procedure first accesses local disk or non-volatile random-access memory (NVRAM) in the node, to read a device driver parameter, referred to herein as MAXLUN. This parameter specifies the maximum number of LUNs the node is allowed to grope; for example a value MAXLUN = 8 allows LUNs 0 to 7 to be groped. Conventionally, MAXLUN would be set to the highest LUN value known to be available. The purpose of this is to reduce system boot time by ensuring that the node does not waste time looking for LUNs that do not exist. However, in the present system, instead of being set to the highest LUN value known to be available, MAXLUN is set to the highest LUN value that the Unix or NT node 11 is to be allowed to access. This ensures that any higher LUNs are not accessible to this node.

The grope procedure then selects each of the HBAs in the node, and sends a SCSI TEST UNIT READY command to each target device (i.e. storage controller) attached to that HBA. If there is no response from a controller within a predetermined timeout period, it moves on to the next controller.

If a valid response is received from a storage controller, the procedure sends a series of SCSI INQUIRY commands to that controller, with LUN values 0 up to MAXLUN. The returned INQUIRY data includes the vendor identity, device type, and a peripheral qualifier. The peripheral qualifier indicates whether the device is "present and currently connected" or "supported but not currently connected". This latter status value is used in dual path systems where a LUN may be owned by either of two storage controllers. The returned data is added to a device table, which is available for use by the operating system.

Figure 3 is a pseudo-code representation of the grope procedure for the ICL VME node 10.

The VME node has a configuration file, stored on a local disk. For each VME controller in the node (essentially an I/O path that maps on to a particular HBA) the configuration file holds an entry, referred to herein as STARTLUN, that tells VME which LUN to start from for that HBA.

The grope procedure selects each HBA, and reads the appropriate STARTLUN from the configuration file. It then sends a SCSI TEST UNIT READY command to each target device (i.e. storage controller) attached to that HBA. If there is no response from a controller within a predetermined timeout period, it moves on to the next controller.

If a valid response is received from a storage controller, the procedure sends a series of SCSI INQUIRY commands to that controller, with LUN values from STARTLUN up to STARTLUN + 31. As before the returned INQUIRY data is added to a device table which is available for use by the operating system.

The grope procedure described above with reference to Figure 3 is similar to the conventional VME grope procedure, except that it is modified to scan upwards from STARTLUN, rather than from zero up to a specified maximum LUN value. This ensures that any lower LUNs are not accessible to this node. The amount of modification of the VME operating system required to achieve this is minimal, and no other special software is required.

In summary, it can be seen that in the system described above, each of the processing nodes 10, 11 has access to a separate set of LUNs, and cannot access the LUNs belonging to the other node. For example, if MAXLUN is set to 8, then the NT/Unix node 11 can access only LUNs 0 to 7, and if STARTLUN is set to 8, then the VME node 10 can access only LUNs 8 to 39.

Some possible modifications

It will be appreciated that many modifications may be made to the system described above without departing from the scope of the present invention.

Although in the above example the system includes a VME node and a Window NT or Unix node, other combinations are possible. In particular, it is possible to include several VME nodes. By suitably choosing the STARTLUN parameters of each node, it is possible to ensure that each of these nodes can access only its own LUNs. For example, if two VME nodes have STARTLUN values of 8 and 40 respectively, the first will recognise LUNs from 8 to 39, and the second from 40 to 71.

In the embodiment described above, the VME node scans LUNs from STARTLUN up to $\text{STARTLUN} + 31$, limiting the number of LUNs to a maximum of 32 per VME controller. However, in an alternative embodiment of the invention, the configuration file may hold an STOPLUN value, in addition to STARTLUN, for each VME controller.

In this case, the grope procedure is arranged to scan from STARTLUN up to STOPLUN.

It will be appreciated that the invention is not restricted to the use of Fibre Channel; for example, conventional SCSI connections may be used.

Instead of RAID arrays, some or all of the logical storage units may comprise stand-alone disks or tape units.

CLAIMS

1. A computer system comprising:
 - at least two processing nodes; and
 - a shared data storage system comprising a number of storage units, each having a logical unit number (LUN) associated with it;
 - wherein a first of the processing nodes stores a maximum LUN value in non-volatile storage, and is arranged to perform, on boot-up, a device discovery procedure in which it scans the shared data storage system to find storage units with LUNs not greater than the maximum LUN value;
 - and wherein a second of the processing nodes stores a start LUN value in non-volatile storage, and is arranged to perform, on boot-up, a device discovery procedure in which it scans the shared data storage system to find storage units with LUNs not less than the start LUN value.
2. A computer system according to Claim 1 wherein the device discovery procedure in the second of the processing nodes scans the shared data storage system to find storage units with LUNs in a range between the start LUN value and the start LUN value plus a predetermined constant value.
3. A computer system according to Claim 1 wherein the second of the processing nodes also stores a stop LUN value in non-volatile storage, and wherein the device discovery procedure in the second of the processing nodes scans the shared data storage system to find storage units with LUNs in a range between the start LUN value and the stop LUN value.
4. A computer system according to any preceding claim wherein the processing nodes are connected to the data storage system by way of a Fibre Channel connection, using the SCSI protocol.



Application No: GB 0002909.0
Claims searched: 1 to 7

Examiner: Julyan Elbro
Date of search: 20 October 2000

Patents Act 1977 Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:
UK CI (Ed.R): G4A (AFL)
Int CI (Ed.7): G06F 9/46, 13/40, 13/42
Other: ONLINE: COMPUTER EPODOC INSPEC JAPIO WPI Selected Internet Sites

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	"Emulex LightPulse 6/7/8000 PCI Fibre Channel Host Adapter SCSIport Information", 14 December 1999, Emulex Corporation, available from http://sup.datadirectnet.com/drivers/emulex/sli124.txt , see especially the registry option "MaximumLun", and the reference to this under "Config Utility".	

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.